

## Chapter 12

# Scientific Discovery and Intelligent Evolution

In previous chapters, we primarily discussed the evolution of agentic systems from a technical perspective, focusing on how to develop systems that can effectively perform well-defined tasks traditionally executed by humans. However, a fundamental and important question remains: can these agents drive a self-sustaining innovation cycle that propels both agent evolution and human progress?

Scientific knowledge discovery is a compelling example of self-evolution in intelligent beings, as it helps them adapt to the world in a sustainable way. Agents capable of discovering scientific knowledge at different levels of autonomy and in a safe manner will also play important roles in technological innovation for humanity. In this section, we survey progress in autonomous discovery using agentic workflows and discuss the technological readiness toward fully autonomous, self-evolving agents. Within this scope, the goal of the agent is to uncover, validate, and integrate data, insights, and principles to advance an objective scientific understanding of natural phenomena. Instead of altering the world, the agent seeks to better understand nature as a Scientist AI [859] and assist humans in extending the boundaries of knowledge.

We first define the concept of knowledge and intelligence to clarify our discussion, then introduce three typical scenarios where agents and scientific knowledge interact. We also highlight existing successes and examples of self-enhancing agents applied to theoretical, computational, and experimental scientific research. Lastly, we summarize the current challenges for a future outlook.

### 12.1 Agent’s Intelligence for Scientific Knowledge Discovery

Knowledge, traditionally defined as *justified true belief*, traces back to Plato [860] and has been further refined by Edmund Gettier [861], who argued that knowledge must be produced by a reliable cognitive process—though its precise definition remains debated [862]. In our discussion, we describe scientific knowledge discovery as the process of collecting data and information to either justify or falsify rational hypotheses about target scientific problems. To discuss the capability of agents in scientific knowledge discovery, we first explore a general framework for measuring an agent’s intelligence through the lens of information theory.

#### 12.1.1 KL Divergence-based Intelligence Measure

**The agent’s intelligence can be measured by the KL divergence between its predicted and real-world probability distributions of unknown information.** A long-standing goal in both artificial intelligence and the philosophy of science is to formalize what it means for an agent to “understand” the world. From Jaynes’ view of probability theory as extended logic for reasoning under uncertainty [863], to Parr et al.’s framing of intelligence as minimizing model-world divergence under the free energy principle [864], many frameworks converge on a common theme: intelligent behavior arises from making accurate predictions about an uncertain world. Clark [344], for instance, argues that intelligent agents constantly engage with the world through prediction and error correction to reduce surprise. Chollet [865] emphasizes that intelligence should reflect skill-acquisition efficiency, because of the dynamic nature of task adaptation. Together, these views suggest that intelligence involves building predictive and adaptable models—an idea formalized here through a probabilistic framework that links reasoning to knowledge acquisition and enables comparison across agents in scientific discovery.

Building on this foundation, we consider intelligence in the specific context of scientific knowledge discovery, where the agent’s primary objective is to infer unknown aspects of the physical world from limited data. From the agent’s perspective in knowledge discovery, the world  $\mathcal{W}$  is characterized by an ensemble of datasets  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  related to the scientific problem the agent aims to understand. During the agent’s interaction with  $\mathcal{W}$ , each dataset appears in the experimental measurements or observations with a probability  $P_{\mathcal{W}}(\mathbf{x})$ . Here we assume that individual data points  $x_i$  may or may not be correlated. For example, in a task of text generation using a language model,  $x_i$  represents a chunk of tokens forming a meaningful proposition, and  $\mathbf{x}$  is a coherent text constructed from known and inferred propositions. In this context, the “world” is the ensemble of all propositions.

Let  $\theta$  denote the parameter that parameterizes the agent’s world model,  $M_t^{\text{wm}}$ , as defined in Table 1.2. For instance, in a transformer model with a fixed architecture,  $\theta$  represents its weights. Given  $\theta$  and a dataset  $\mathbf{x}$ , the agent predicts a probability distribution  $P_{\theta}(\mathbf{x})$ . In general, different AI agents could be optimized for different goals. For scientific knowledge discovery, we assume that the agent’s goal is to produce a good description of the real world, i.e., a world model that predicts yet-to-be-explored natural phenomena as accurately as possible. A more intelligent agent produces a better approximation of the real-world distribution  $P_{\mathcal{W}}(\mathbf{x})$ . The agent’s intelligence can thus be measured by the KL divergence, or relative entropy, between these two probability distributions:

$$D_0(\theta) = \sum_{\mathbf{x} \subseteq \mathcal{W}} P_{\mathcal{W}}(\mathbf{x}) \log \frac{P_{\mathcal{W}}(\mathbf{x})}{P_{\theta}(\mathbf{x})} \quad (12.1)$$

$D_0(\theta)$  describes the difference between  $P_{\mathcal{W}}(\mathbf{x})$  and  $P_{\theta}(\mathbf{x})$ . More precisely, in the context of hypothesis testing, if we sample  $P_{\mathcal{W}}(\mathbf{x})$   $N$  times and compare the results with the predictions from  $P_{\theta}(\mathbf{x})$ , the probability of mistaking  $P_{\mathcal{W}}(\mathbf{x})$  for  $P_{\theta}(\mathbf{x})$  scales as  $e^{-ND_0(\theta)}$  [866]. In other words, an agent with a lower  $D_0(\theta)$  produces predictions that align more closely with reality.

For example, consider two materials synthesis agents whose goal,  $M_t^{\text{goal}}$ , is to understand whether or not an inorganic compound of interest,  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$ , is synthesizable. The agents can predict either (1)  $\mathbf{x}_1 = \{\text{CaFe}_2(\text{PO}_4)_2\text{O} \text{ is synthesizable}\}$ , and (2)  $\mathbf{x}_2 = \{\text{CaFe}_2(\text{PO}_4)_2\text{O} \text{ is not synthesizable}\}$ . In reality, since  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$  is a natural mineral,  $P_{\mathcal{W}}(\mathbf{x}_1) = 1$  and  $P_{\mathcal{W}}(\mathbf{x}_2) = 0$ . However, this mineral was only recently reported on October 4, 2023[ref], after the knowledge cutoff of many LLMs; thus, the agents lack that knowledge. Compare Agent 1, which guesses randomly  $P_{\theta_1}(\mathbf{x}_1) = P_{\theta_1}(\mathbf{x}_2) = 0.5$ , yielding  $D_0(\theta_1) = \log 2$ . In contrast, Agent 2 uses first-principles calculations and finds that  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$  (assume structure is xx [cite: Materials Project ID]) is the lowest-energy phase among its competitors [ref], indicating stability. Thereby, Agent 2 predicts that  $\text{CaFe}_2(\text{PO}_4)_2\text{O}$  is likely synthesizable, suggesting  $P_{\theta_2}(\mathbf{x}_1) > 0.5 > P_{\theta_2}(\mathbf{x}_2)$ . Consequently,  $D_0(\theta_2) = -\log P_{\theta_2}(\mathbf{x}_1) < D_0(\theta_1)$ , meaning that Agent 2 has a more accurate understanding of the real world.

Now, let us assume the agent has conducted some measurements and determined specific values for a subset of data points  $x_i$ . Let  $\mathbf{x}_K$  denote this known subset and  $\mathbf{x}_U$  the remaining unknown part. Correspondingly, we define the space of all existing knowledge as  $\mathcal{K}$  and the space of all unknown information as  $\mathcal{U}$ , satisfying  $\mathbf{x}_K \subseteq \mathcal{K}$ ,  $\mathbf{x}_U \subseteq \mathcal{U}$ , and  $\mathcal{K} \cup \mathcal{U} = \mathcal{W}$ . For example, in text generation, the the prompt text  $\mathbf{x}_K$  represents already known information. The efficiency of the language model is then measured by its predictive accuracy for the generated text  $\mathbf{x}_U$  based on  $\mathbf{x}_K$ . More generally, the agent’s intelligence is measured by the relative entropy of the conditional probability distribution:

$$D_K(\theta, \mathbf{x}_K) = \sum_{\mathbf{x} \subseteq \mathcal{U}} P_{\mathcal{W}}(\mathbf{x}|\mathbf{x}_K) \log \frac{P_{\mathcal{W}}(\mathbf{x}|\mathbf{x}_K)}{P_{\theta}(\mathbf{x}|\mathbf{x}_K)} \quad (12.2)$$

In practice, all of the agent’s knowledge is stored in its memory  $M_t^{\text{mem}}$ , i.e.,  $\mathbf{x}_K = \mathcal{K} = M_t^{\text{mem}}$  and  $\mathcal{U} = \mathcal{W} \setminus M_t^{\text{mem}}$ , we define the agent’s intelligence as:

$$IQ_t^{\text{agent}} \equiv -D_K(\theta, M_t^{\text{mem}}) = - \sum_{\mathbf{x} \subseteq \mathcal{U}} P_{\mathcal{W}}(\mathbf{x}|M_t^{\text{mem}}) \log \frac{P_{\mathcal{W}}(\mathbf{x}|M_t^{\text{mem}})}{P_{\theta}(\mathbf{x}|M_t^{\text{mem}})} \quad (12.3)$$

In other words, the the agent’s intelligence  $IQ_t^{\text{agent}}$  is determined by its memory  $M_t^{\text{mem}}$  and the parameter  $\theta$  of its world model  $M_t^{\text{wm}}$ . A schematic plot is shown in Figure 12.1. At time  $t = 0$ , when the  $M_t^{\text{mem}}$  is very limited or lack relevant information to a new target scientific problem,  $IQ_t^{\text{agent}}$  is primarily determined by the zero-shot predictive ability of  $M_t^{\text{wm}}$ , corresponding to fluid intelligence [867]. Over time, as more relevant knowledge is incorporated into  $M_t^{\text{mem}}$ ,  $IQ_t^{\text{agent}}$  becomes increasingly dependent on the knowledge-augmented predictive capability of  $M_t^{\text{wm}}$ , reflecting crystallized intelligence [868].

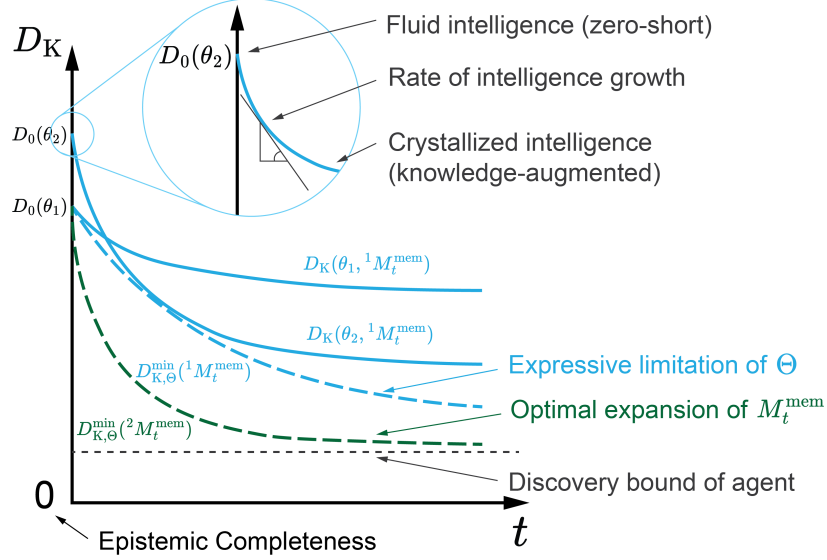


Figure 12.1: **Schematic representation of agent intelligence and knowledge discovery.** The agent’s intelligence, measured by the KL divergence  $D_K$  between predictions and real-world probability distributions, evolves from fluid intelligence (zero-shot predictions for new problems) to crystallized intelligence (knowledge-augmented predictions after learning) as it accumulates data in its memory  $M_t^{\text{mem}}$  over time  $t$ . Given  $M_t^{\text{mem}}$ , the evolution of  $D_K$  varies within the world model’s parameter space  $\Theta$ , as illustrated by  $\theta_1$  and  $\theta_2$  in the solid lines. The expressive limitation of  $\Theta$  is characterized by the envelope  $D_{K,\Theta}^{\min}$ . Given  $\Theta$ ,  $D_{K,\Theta}^{\min}$  is influenced by different knowledge expansion strategies, such as  ${}^1M_t^{\text{mem}}$  and  ${}^2M_t^{\text{mem}}$ , shown as dash lines.

### 12.1.2 Statistical Nature of Intelligence Growth

**The agent’s intelligence, in a statistical sense, is a non-decreasing function of acquired knowledge.** Roughly speaking,  $IQ_t^{\text{agent}}$  quantifies both the amount of knowledge an agent has acquired and how effectively the agent can apply that knowledge after learning from  $M_t^{\text{mem}}$ . Intuitively, if the agent gains additional information at time  $t$ —which corresponds to enlarging  $M_t^{\text{mem}}$  and shrinking  $\mathcal{U}$ —its intelligence should increase.

To understand this process, consider a small region  $\Delta \subseteq \mathcal{U}$  and examine the effect of adding a dataset  $\mathbf{x}_\Delta$  from  $\Delta$  to  $M_t^{\text{mem}}$ . Denote  $\mathcal{U} = \mathcal{U}' \cup \Delta$ , where  $\mathcal{U}'$  represents the remaining unknown part of the world. The agent’s intelligence at time  $t + 1$  is given by:

$$IQ_{t+1}^{\text{agent}} \equiv -D_K(\theta, M_t^{\text{mem}} \mathbf{x}_\Delta) = - \sum_{\mathbf{x}' \subseteq \mathcal{U}'} P_{\mathcal{W}}(\mathbf{x}' | M_t^{\text{mem}} \mathbf{x}_\Delta) \log \frac{P_{\mathcal{W}}(\mathbf{x}' | M_t^{\text{mem}} \mathbf{x}_\Delta)}{P_\theta(\mathbf{x}' | M_t^{\text{mem}} \mathbf{x}_\Delta)} \quad (12.4)$$

Directly comparing  $IQ_t^{\text{agent}}$  and  $IQ_{t+1}^{\text{agent}}$  is challenging. Instead, we can compare the expected value of  $IQ_{t+1}^{\text{agent}}$ , averaging over  $\mathbf{x}_\Delta$  with probability  $P_{\mathcal{W}}(\mathbf{x}_\Delta | M_t^{\text{mem}})$ . This expectation represents the average amount of knowledge gained by measuring  $\Delta$ , given prior knowledge in  $M_t^{\text{mem}}$ . We obtain:

$$\begin{aligned} \sum_{\mathbf{x} \subseteq \Delta} P_{\mathcal{W}}(\mathbf{x} | M_t^{\text{mem}}) IQ_{t+1}^{\text{agent}} &= - \sum_{\mathbf{x}' \subseteq \mathcal{U}', \mathbf{x} \subseteq \Delta} P_{\mathcal{W}}(\mathbf{x}' \mathbf{x} | M_t^{\text{mem}}) \log \frac{P_{\mathcal{W}}(\mathbf{x}' | M_t^{\text{mem}} \mathbf{x})}{P_\theta(\mathbf{x}' | M_t^{\text{mem}} \mathbf{x})} \\ &= IQ_t^{\text{agent}} + \sum_{\mathbf{x} \subseteq \Delta} P_{\mathcal{W}}(\mathbf{x} | M_t^{\text{mem}}) \log \frac{P_{\mathcal{W}}(\mathbf{x} | M_t^{\text{mem}})}{P_\theta(\mathbf{x} | M_t^{\text{mem}})} \end{aligned} \quad (12.5)$$

The second term is the relative entropy of the conditional probability distribution of  $\mathbf{x}_\Delta$  conditioned on  $M_t^{\text{mem}}$ , which is always non-negative. Therefore, on average,  $IQ_{t+1}^{\text{agent}}$  is non-decreasing as  $M_t^{\text{mem}}$  acquires new knowledge over time. Note that  $IQ_{t+1}^{\text{agent}}$  can be further increased by leveraging the newly acquired knowledge to optimize  $\theta$  within  $M_t^{\text{wm}}$ .

Interestingly, the expected gain in intelligence at time  $t$  is determined by the discrepancy between the actual distribution  $P_{\mathcal{W}}(\mathbf{x} | M_t^{\text{mem}})$  and the model-predicted distribution  $P_\theta(\mathbf{x} | M_t^{\text{mem}})$ . In other words, the rate of intelligence growth in Figure 12.1 is higher when the new measurement result is more unexpected. This observation identifies scientist agents

[859] as a special type of curiosity-driven agent [869], prioritizing exploration over exploitation to expand the frontiers of knowledge for deeper understanding of nature. Unlike agents that leverage existing knowledge to achieve predefined objectives, curiosity-driven agents can learn without extrinsic rewards [387, 870] (see Section 5.3 for details), enabling discoveries beyond human-planned search spaces and revealing knowledge in unexplored domains. This potential also underscores the importance of equipping curiosity-driven agents with fundamental perception and action tools that can be transferred to explore new knowledge domains.

### 12.1.3 Intelligence Evolution Strategies

**The strategy for expanding known information determines how quickly an agent’s intelligence evolves.** For a given knowledge base  $M_t^{\text{mem}}$ , the parameter  $\theta$  can be optimized over a space of world models  $\Theta$  characterized by the architecture of  $M_t^{\text{wm}}$ . The optimal agent is the one that minimizes  $D_K(\theta, M_t^{\text{mem}})$ , thereby maximizing  $IQ_t^{\text{agent}}$ :

$$\theta_{K,t}^* \equiv \arg \sup_{\theta} IQ_t^{\text{agent}} = \arg \inf_{\theta} D_K(\theta, M_t^{\text{mem}}) \quad (12.6)$$

and

$$D_{K,\Theta}^{\min}(M_t^{\text{mem}}) \equiv D_K(\theta_{K,t}^*, M_t^{\text{mem}}) \quad (12.7)$$

Here,  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$  represents the minimum unknown after learning from  $M_t^{\text{mem}}$  for this family of models, quantifying the expressive limitations of  $\Theta$ . As shown in Figure 12.1,  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$  forms the envelope of the family of functions  $D_K(\theta, M_t^{\text{mem}})$ , where  $\theta$  ranges over  $\Theta$ .

For a given model family  $\Theta$ ,  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$  measures the best possible prediction of residual unknowns in addressing the target scientific problem based on  $M_t^{\text{mem}}$ . In other words, the knowledge content in  $M_t^{\text{mem}}$  is captured by  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$ . One can prove that  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$  is monotonically non-increasing as  $M_t^{\text{mem}}$  expands, since it forms the envelope of a family of non-increasing functions  $D_K(\theta, M_t^{\text{mem}})$ . This expansion process is tied to how the agent acts and gains information, driven by  $M_t^{\text{wm}}$ , which determines the optimal expansion and executes it through the action  $a_t \in \mathcal{A}$  at time  $t$  (see Table 1.2).

During knowledge discovery, different strategies can be employed to expand  $M_t^{\text{mem}}$ . The optimal expansion strategy is the one that results in the steepest decrease of  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$ . For instance, in Figure 12.1, we illustrate two strategies for expanding  $M_t^{\text{mem}}$ , denoted as  $^1M_t^{\text{mem}}$  and  $^2M_t^{\text{mem}}$ . The first strategy,  $^1M_t^{\text{mem}}$ , represents random exploration, while the second,  $^2M_t^{\text{mem}}$ , follows a hypothesis-driven approach [871] in which the agent first formulates a hypothesis about the underlying mechanism of the target problem and then designs an experiment to justify or falsify this hypothesis [749]. In practice, experimentalists typically adopt the hypothesis-driven strategy because it enables them to guide the expansion of  $M_t^{\text{mem}}$  in a way that maximizes the reduction of  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$ , subject to resource constraints. This approach is generally more efficient than random exploration for expanding  $M_t^{\text{mem}}$ , leading to  $D_{K,\Theta}^{\min}(^2M_t^{\text{mem}})$  descending faster than  $D_{K,\Theta}^{\min}(^1M_t^{\text{mem}})$ .

In general, the knowledge discovery process proceeds iteratively, repeatedly optimizing the world model parameter  $\theta$  to approach  $\theta_{K,t}^*$  and expanding  $M_t^{\text{mem}}$  in a rational manner to accelerate the decrease of  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$ . The ideal state is achieving epistemic completeness, i.e.,  $D_{K,\Theta}^{\min}(M_t^{\text{mem}}) = 0$ , meaning zero discrepancy between the agent’s prediction and the real-world phenomena. However, for a specific agent, a discovery bound may exist, where  $D_{K,\Theta}^{\min}(M_t^{\text{mem}})$  approaches zero but remains positive. These discrepancies arise from practical constraints and the limitations of  $\Theta$ ,  $\mathcal{A}$ , and other design spaces of the agent [872]. Achieving a low discovery bound requires designing an adaptive world model architecture, an efficient knowledge expansion strategy, and a sufficient action space.

## 12.2 Agent-Knowledge Interactions

Typical forms of scientific knowledge include observational knowledge (e.g., experimental measurements, computational results), methodological knowledge (e.g., experimental methods, computational techniques, protocols), and theoretical knowledge (e.g., theories, laws, predictive models). These forms of knowledge can contribute to scientific understanding as long as they consist of data and information processed in a way that affects the probability distribution of unknown information  $P_{\theta}(\mathbf{x}_U | M_t^{\text{mem}})$ , reduces  $D_K(\theta, M_t^{\text{mem}})$ , and facilitates decision-making.

In principle, external scientific knowledge has been shown to be useful in improving agent performance in reasoning and decision-making [873, 874]. However, the scope of this survey lies in how agents can autonomously discover and utilize knowledge to enhance themselves. Scientific knowledge discovery workflows typically involve hypothesis generation, protocol planning, conducting experiments and computations, analyzing data, deriving implications, and

revising hypotheses—often as part of an iterative cycle. An agent that can perceive, learn, reason, and act has the potential to drive such workflows in an autonomous manner, for example by using application programming interfaces (APIs) to interact with physical instruments to acquire scientific knowledge and iteratively enhance its knowledge base (Figure 12.2). The agent will use the acquired knowledge to update its mental states  $M_t$  to make better decisions when interacting with the world  $\mathcal{W}$ . We will now highlight three scenarios where agents discover scientific knowledge and enhance themselves.

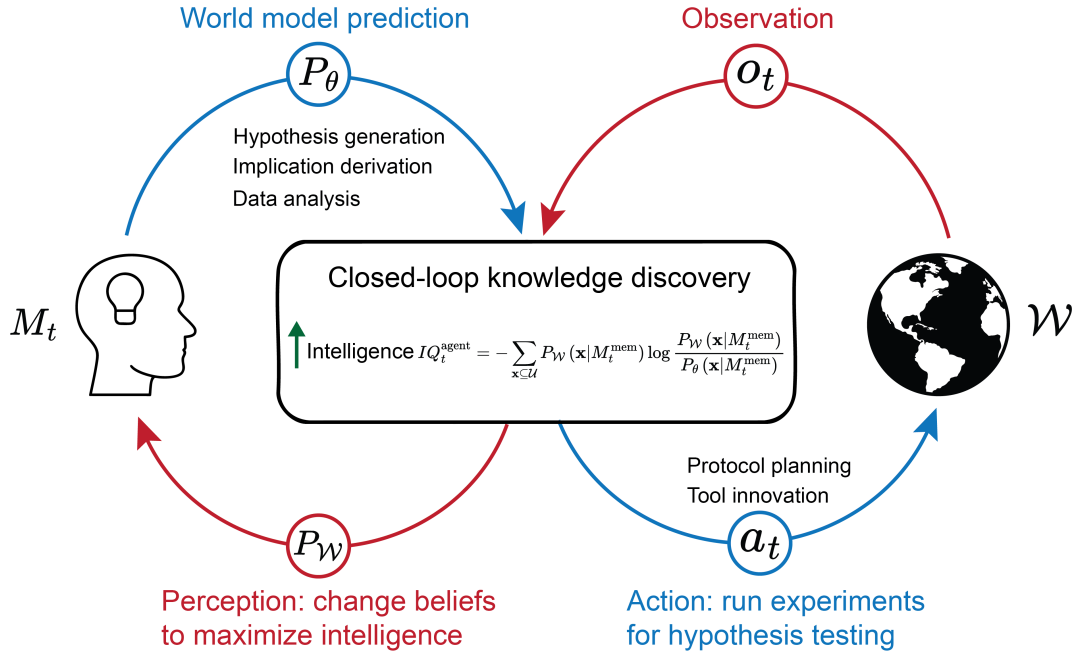


Figure 12.2: **Closed-loop knowledge discovery for sustainable self-evolution.** The agent aims to iteratively enhance its intelligence  $I Q_t^{agent}$  through hypothesis generation and testing, as well as through data analysis and implication derivation. When interacting with the physical world  $\mathcal{W}$ , the agent generates hypotheses as an explicitly or implicitly predicted distribution ( $P_{\theta}$ ) of unknown information, takes actions ( $a_t$ ) for hypothesis testing, observes experimental results ( $o_t$ ), and updates beliefs based on perception of the real-world distribution ( $P_{\mathcal{W}}$ ). When not interacting with  $\mathcal{W}$ , the agent distills knowledge from existing data and premises, updating mental states  $M_t$  directly. Inspired by Figures 2.3 and 2.5 in [864].

### 12.2.1 Hypothesis Generation and Testing

Hypothesis generation and testing (Figure 12.2) is a critical application of agents in autonomous scientific discovery, as it has the potential to enable outside-the-box innovations [749]. In essence, hypothesis generation is the formation of potential rules that govern data distribution—ranging from single observations to large datasets—pertaining to unobserved scientific phenomena. According to Sir Karl Popper, a scientific hypothesis must be falsifiable [875, 876]; in this discussion, we define a hypothesis that survives falsification as a *justified true hypothesis* [877, 860]. Typically, scientists test hypotheses by conducting experiments to either justify or falsify them. A hypothesis is considered more valuable if it is broad enough to explain a wide range of data and is highly likely to be true.

To tackle a scientific problem, the agent formulates one or a small number of high-value hypotheses based on its mental state  $M_t$ , which contains only incomplete information about the partially observable world  $\mathcal{W}$ . After testing through experiments or computations, a *justified true hypothesis* becomes instructive knowledge, expanding  $M_t^{mem}$  in a way that rapidly minimizes  $D_{K,\Theta}^{min}(M_t^{mem})$ . Hence, generating and testing high-value hypotheses can quickly promote knowledge discovery and increase  $I Q_t^{agent}$ . In this scenario, the agent employs the learning function,  $L$ , to process observations from hypothesis testing,  $o_t$ , into knowledge and update its mental states  $M_t$ .

**Generating physically meaningful hypotheses is a key step.** The agent typically uses LLMs along with collaborative architectures and domain knowledge for hypothesis generation [878]. Si et al. [742] conducted a large-scale human study involving over 100 NLP researchers, and found that LLM-generated ideas were rated as more novel ( $p < 0.05$ ) than human expert ideas, albeit slightly weaker in feasibility. Ghafarollahi et al. [743] developed SciAgents, which generates



and refines materials science hypotheses to elucidate underlying mechanisms, design principles, and unexpected properties of biologically inspired materials. Based on large-scale ontological knowledge graphs, SciAgents samples a viable path between concepts of interest, formulates a pertinent hypothesis, and expands it into a full research proposal with detailed hypothesis-testing methods and criteria. It employs two dedicated agents to review, critique, and improve the proposed hypothesis, but does not include the step of hypothesis testing through actual experiments. Similarly, Su et al. [879] and Baek et al. [880] proposed leveraging teamwork—such as collaborative discussions and agent critics—to produce novel and effective scientific hypotheses. In addition, Gower et al. [881] introduced LGEM<sup>+</sup>, which utilizes a first-order logic framework to describe biochemical pathways and generate 2,094 unique candidate hypotheses for the automated abductive improvement of genome-scale metabolic models in the yeast *S. cerevisiae*.

#### **Hypotheses only become knowledge after being justified through computational or experimental observations.**

Lu et al. [745] introduced the AI Scientist, a system designed for fully automated scientific discovery. The AI Scientist can conduct research independently and communicate its findings, as demonstrated in three machine learning subfields—diffusion modeling, transformer-based language modeling, and learning dynamics. It generates original research ideas, writes code, performs computational experiments, visualizes results, drafts complete scientific papers, and even simulates a peer review process for evaluation. For instance, it proposed the hypothesis that “adaptive dual-scale denoising can improve diffusion models by balancing global structure and local details in generated samples,” which was justified through image generation tests on four 2D datasets. Similarly, Schmidgall et al. [746] developed the Agent Laboratory to autonomously carry out the entire research process, including literature review, computational experimentation, and report writing. They evaluated Agent Laboratory’s capability for knowledge discovery by addressing five research questions in computer vision and natural language processing, achieving an average human-evaluated experiment quality score of 3.2 out of 5. In addition, Tiukova et al. [744] developed Genesis, an automated system capable of controlling one thousand  $\mu$ -bioreactors, performing mass spectrometry characterization, accessing a structured domain information database, and applying experimental observations to improve systems biology models. Genesis can initiate and execute 1,000 hypothesis-driven closed-loop experimental cycles per day. Using a similar approach, the Genesis team has advanced the yeast (*S. cerevisiae*) diauxic shift model, outperforming the previous best and expanding its knowledge by 92 genes (+45%) and 1,048 interactions (+147%) [882]. This knowledge also advances our understanding of cancer, the immune system, and aging. Similarly, Gottweis et al. [749] introduced the AI co-scientist, which autonomously generates and refines novel research hypotheses, with *in vitro* validation in three biomedical areas: drug repurposing, novel target discovery, and mechanisms of bacterial evolution and antimicrobial resistance.

**Discovered knowledge enhances the agent’s mental states, such as  $M_t^{\text{mem}}$ ,  $M_t^{\text{wm}}$ , and  $M_t^{\text{rew}}$ .** Tang et al. [747] developed ChemAgent, which improves chemical reasoning through a dynamic, self-updating memory,  $M_t^{\text{mem}}$ . ChemAgent proposes hypothetical answers to chemistry questions in a development dataset, evaluates them against the ground truth, and simulates the hypothesis-testing process used in real-world research. Correct answers are then stored as knowledge in its memory to support future chemistry question answering. This self-updating memory resulted in performance gains of up to 46% (with GPT-4) when ChemAgent was applied to four chemical reasoning datasets from SciBench [883]. Wang et al. [884] introduced Molecular Language-Enhanced Evolutionary Optimization (MOLLEO), which iteratively proposes hypotheses for modifying candidate drug molecules in  $M_t^{\text{mem}}$ , evaluates their drug-likeness and activity, and updates the candidates in  $M_t^{\text{mem}}$  to enhance drug discovery. Similarly, Jia et al. [885] developed LLMatDesign, which employs hypothesis-guided structure generation and a self-updating  $M_t^{\text{mem}}$  to design inorganic photovoltaic materials, whose ideality is defined by matching the target band gap and having the most negative formation energy.

Sim et al. [748] introduced ChemOS 2.0, which orchestrates closed-loop operations in chemical self-driving laboratories (SDLs). ChemOS 2.0 integrates *ab initio* calculations, experimental orchestration, and statistical algorithms for the autonomous discovery of high-performance materials. A case study on discovering organic laser molecules demonstrates its capabilities. It employs a Bayesian optimizer, Altas, as its world model  $M_t^{\text{wm}}$  to predict the optical properties of hypothetical molecules—specifically Bis[(N-carbazole)styryl]biphenyl (BSBCz) derivatives—including gain cross section and spectral grain factor. Based on these predictions, ChemOS 2.0 recommends molecules with a higher probability of success in the experimental campaign. It then utilizes an optical characterization platform and the AiiDA software package to measure and simulate the properties of test molecules. The results are used to update  $M_t^{\text{wm}}$ , improving the accuracy of future experimental predictions.

Hysmith et al. [886] published a perspective highlighting the crucial role of reward function design in developing forward-looking workflows for SDLs. Agents can be highly effective at solving POMDP problems in simulated environments, such as computer games or simulations, but often struggle with real-world applications. A well-defined reward function is essential for iterative self-evolution. However, in many real-world scientific research problems, reward functions are ill-defined or absent at the end of experimental campaigns due to the lack of direct measurements,

the complexity of experimental results, and the need to balance multiple objectives. The discovery of new knowledge can serve as a valuable resource for refining  $M_t^{\text{rew}}$ , guiding hypothesis exploration and experimental data collection.

### 12.2.2 Protocol Planning and Tool Innovation

The capability to plan experimental protocols and optimize tool usage enables the agent to solve complex scientific puzzles within the autonomous discovery loop. As introduced in Section 9.4, the agent can systematically evaluate and refine its approach to selecting, invoking, and integrating available tools—and even develop new tools tailored to specific task requirements. While optimized protocols and tool usage do not directly reduce  $D_K(\theta, M_t^{\text{mem}})$ , they enhance execution efficiency and effectiveness in refining the probability distribution of unknown information,  $P_\theta(\mathbf{x}_U | M_t^{\text{mem}})$ , thereby accelerating knowledge discovery. In this scenario, the agent leverages the reasoning function  $R$  to translate its evolving mental states  $M_t$ , continuously updated with new knowledge, into real-world actions  $a_t$  for more effective and faster hypothesis testing (Figure 12.2).

**Scheduling and orchestrating the selection and recombination of existing tools is critical.** Scientific experiments typically depend on diverse instruments for analyzing reaction products, with decisions rarely rely on just one measurement. Effectively utilizing necessary instruments without wasting resources and time requires the agent to learn to use tools in an integrated and adaptive manner. Dai et al. [750] designed a modular workflow that integrates mobile robots, an automated synthesis platform, and various characterization instruments for autonomous discovery. They exemplified this system across three domains: structural diversification chemistry, supramolecular host-guest chemistry, and photochemical synthesis. The mobile robot follows a synthesis-analysis-decision cycle to mimic human experimental strategies, autonomously determining subsequent workflow steps. It selects appropriate instruments, such as the Chemspeed ISynth platform for synthesis, a liquid chromatography-mass spectrometer (UPLC-MS) for measuring mass spectra corresponding to chemical peak signals, and a benchtop nuclear magnetic resonance spectrometer (NMR) for tracking chemical transformations from starting materials to products.

Beyond individual laboratories, tool orchestration is essential for delocalized and asynchronous scientific discovery. Strieth-Kalthoff et al. [751] demonstrated a closed-loop integration of five materials science laboratories across three continents, advancing delocalized and democratized scientific discovery. These five laboratories have varying strengths—for example, the University of British Columbia specializes in continuous preferential crystallization, while Kyushu University excels in thin film fabrication and characterization. Strieth-Kalthoff et al. employed a cloud-based experiment planner to continuously learn from the incoming data and effectively prioritize informative experiments across the five laboratories, resulting in the discovery of 21 new state-of-the-art materials for organic solid-state lasers.

**Moreover, the agent can optimize existing tools and even create new ones to enhance its capabilities.** Swanson et al. [752] developed the Virtual Lab, an AI-driven research environment that facilitated the design and experimental validation of new SARS-CoV-2 nanobodies. Within the Virtual Lab, AI agents conduct scientific discussion in team meetings and execute specialized tasks in individual sessions. One key agenda for the agents was developing tools to aid in the design of nanobody binders [887], including: (1) a sequence analysis tool that ranks candidate point mutations using log-likelihood ratios from the ESM protein language model [888]; (2) a structure evaluation tool that extracts interface pLDDT scores from AlphaFold-Multimer predictions [889], offering a proxy for antibody-antigen binding affinity; and (3) an energy estimation tool built on Rosetta [890] to quantify binding strength between nanobody variants and the spike protein. These agent-generated tools enabled the Virtual Lab to discover two novel nanobodies with enhanced binding to the JN.1 or KP.3 SARS-CoV-2 variants, while preserving strong affinity for the ancestral viral spike protein.

### 12.2.3 Data Analysis and Implication Derivation

Although most knowledge discovery processes rely on generating hypotheses and testing them in the real world—where observations  $o_t$  are essential—a significant portion of knowledge can be derived purely through internal actions such as iterative reasoning and deep thinking, which are common in theoretical disciplines. For example, all theorems in Euclidean geometry can be deduced from just five axioms, but these theorems do not explicitly exist in the mental state before they are derived. Given all necessary premises, such as Euclid’s five postulates, the true probability of a hypothesis may remain elusive. However, using deductive and inductive reasoning to draw implications from known premises and data can help either justify or falsify hypotheses, thus reducing  $D_K(\theta, M_t^{\text{mem}})$  and enhancing  $IQ_t^{\text{agent}}$  (Figure 12.2). In this scenario, the agent employs the cognition function  $C$  to use prior mental states  $M_{t-1}$  and internal actions  $a_t$  to derive new knowledge and update mental states to  $M_t$ .

**Deductive reasoning enables knowledge derivation through logic.** Trinh et al. [753] developed AlphaGeometry for the forward deduction of new mathematical theorems based on existing theorems in Euclidean plane geometry. AlphaGeometry employs a neural language model to construct auxiliary points in plane geometry problems and

integrates specialized symbolic engines to exhaustively deduce new true statements, thereby expanding the joint closure of known truths. By leveraging this expanded closure, it alternates between auxiliary constructions and symbolic reasoning engines to uncover further implications. AlphaGeometry demonstrated remarkable performance on a test set of 30 recent Olympiad-level problems, solving 25—more than double the 10 problems solved by the previous best method—and coming close to the level of an average International Mathematical Olympiad (IMO) gold medalist.

**Inductive reasoning enables knowledge derivation through pattern recognition and statistical learning.** Liu et al. [754] introduced the Team of AI-made Scientists (TAIS) to simulate the role of a data scientist for streamlined data analysis. TAIS decomposes a complex data analysis problem into different computational tasks, including coding, self-critique, and regression analysis, to extract meaningful insights from complex datasets. When applied to identifying disease-predictive genes, TAIS achieved an overall success rate of 45.73% on a benchmark dataset containing 457 genetic questions. Ideally, the extracted insights should be logically sound; otherwise, they must be discarded to ensure only accurate findings are safely integrated into mental states. However, limitation in data coverage and the implementation of analysis algorithms may lead to hallucinated insights, underscoring the need for reliable data analyzers and reasoning tools to prevent over-analysis.

## 12.3 Technological Readiness and Challenges

The self-evolution of agents, which in turn drives the advancement of human knowledge, is promised by their early success in the innovation cycle. This cycle involves generating meaningful hypotheses, designing real-time testing protocols, coordinating various experimental and computational tools, analyzing data, deriving implications, and engaging in self-reflection. However, achieving fully autonomous self-evolution remains a significant challenge, given the current technology readiness levels (TRLs) of three fundamental capabilities: real-world interaction, complex reasoning, and the integration of prior knowledge. Further technological progress is required to improve the cycle of self-driven innovation.

### 12.3.1 Real-World Interaction Challenges

Agents interact with the real world primarily through application programming interfaces (APIs). While numerous demonstrations [891] have shown their strong capability to use various APIs, a significant bottleneck in autonomous knowledge discovery remains: the lack of APIs that allow agents to directly execute tasks in a physical laboratory. Physical APIs—interfaces that enable direct control of lab equipment—are far less abundant than computational APIs due to the significant investment of time, expertise, and cost required to develop them. Although existing autonomous laboratories have shown promise, they remain in an early developmental stage (typically TRL 4–6), where straightforward replication or scale-up is challenging. Consequently, building further systems or broadening their application across additional scientific domains still requires substantial customization to address domain-specific needs, along with specialized expertise.

Two key tasks are essential for enabling real-world interaction: *operating lab devices* and *transferring samples between devices*. Seamless integration of physical hardware and experimental samples is crucial to maintaining uninterrupted workflows. However, most experimental instruments are originally designed for human operation. Making them accessible to agents requires extensive efforts across multiple disciplines, including robotics, electrical engineering, mechanical engineering, and software programming. The rising prominence of SDLs is catalyzing the transformation of human-operated devices into agent-accessible systems through APIs. In autonomous labs conducting complex experiments, two parallel and often complementary approaches are commonly adopted to integrate hardware with agentic systems. Both approaches are modular, reconfigurable, and valuable, yet they require ongoing, dedicated development.

**Approach 1: API Integration via Direct Device Adaptation.** This approach involves equipping individual devices with dedicated mechanical adaptations and I/O controllers, enabling them to receive and execute commands from a central control PC. For example, to achieve solid-state synthesis and structural characterization of inorganic materials, A-lab has implemented 16 types of devices to automate experimental tasks such as powder dosing, heating, and diffraction [892]. This approach allows laboratories to function as fully integrated entities by maximizing device utilization, optimizing space and resources, and enabling bespoke tools. However, it is costly, time-consuming, and requires expert knowledge to prototype or retrofit devices for automation. Large language models (LLMs) have been applied to facilitate access to diverse tools, as illustrated by CACTUS, a Chemistry Agent Connecting Tool-Usage to Science [893].

A more accessible alternative for small teams is the *cloud lab* or *science factory* [894], where responsibility for device engineering shifts from individual laboratories to dedicated user facilities or commercial service providers. For



instance, Boiko et al. [895] demonstrated an autonomous chemical research agent, Coscientist, capable of carrying out cross-coupling Suzuki and Sonogashira reactions using experimental setups at the Emerald Cloud Lab [896]. However, cloud labs offer only a fixed set of pre-built devices optimized for common procedures, posing potential challenges for researchers whose experiments require equipment customization, as integrating non-standard tools may involve a lengthy process of negotiation and development.

**Approach 2: Robotic Operation of Experimental Devices.** This approach involves using mobile robots or robotic arms to operate existing devices and transfer samples. In many cases, robots can interact with instruments without modification, apart from minor adjustments such as adding specialized actuators, grippers, or holders. For example, Dai et al. [750] employed mobile robots to explore synthetic chemistry. In their autonomous laboratory, mobile robots enable physical linkages between synthesis and analysis devices that are spatially separated, automating sample transportation and handling. In principle, the robots can perform all actions human researchers require in the laboratory. However, current robotic systems still rely on human pre-programming to map the lab layout, define movement trajectories, and register device positions. Handling unexpected or adaptive situations remains a challenge, as pre-programming cannot anticipate every possible state of an experimental setup. Real-time learning and adaptive manipulation are active areas of research that require further technological advancements. In the long term, embodied AI [897] is expected to enhance robotic learning, allowing agents to quickly adapt to new environments and tools.

The two approaches can be combined. For example, Vescovi et al. [894] define a modular laboratory robotics architecture that allows for translating high-level commands into specific operations for a variety of different robotic apparatus and laboratory equipment, and for linking robotic apparatus with other elements of an AI-driven discovery architecture, such as high-performance computing [898]. This architecture has been used to automate experiments in both the biological and physical sciences [899]. Similarly, Fernando et al. [900] integrate a Robotic Operating System 2 (ROS2) compatible robot into the Bluesky experimental orchestration framework. Lo et al. [901] argue for the development and integration of low-cost “frugal twins” of more expensive equipment to facilitate experimentation and democratize access.

### 12.3.2 Complex Reasoning Challenges

A fundamental philosophical question is whether agents, often powered by LLMs, can truly perform reasoning. By definition, language models generate outputs by predicting the next token, a mechanism fundamentally different from human reasoning. From an outcome-driven perspective, these input-output systems exhibit reasoning ability phenomenologically, as they produce meaningful outputs compared to a reference system generating arbitrary responses [902]. However, regardless of the perspective taken, this capability remains imperfect—particularly when handling complex logical and numerical problems, which are crucial for scientific knowledge discovery.

**Agents and LLMs struggle with hard reasoning tasks.** Glazer et al. [903] introduced FrontierMath, a benchmark comprising hundreds of original and challenging mathematics problems covering most major branches of modern mathematics. Evaluation of state-of-the-art LLM-driven agents—including o1-preview (OpenAI), o1-mini (OpenAI), GPT-4o (OpenAI, 2024-08-06 version), Claude 3.5 Sonnet (Anthropic, 2024-10-22 version), Grok 2 Beta (XAI), and Gemini 1.5 Pro 002 (Google DeepMind)—revealed that no model achieved even a 2% success rate on the full benchmark. Chen et al. [873] presented ScienceAgentBench, a benchmark designed to evaluate language agents in data-driven scientific discovery. Among 102 tasks derived from 44 peer-reviewed publications across four disciplines, OpenAI o1 successfully solved only 42.2% of them. Chollet [865] proposed the Abstraction and Reasoning Challenge (ARC) to assess LLMs’ ability to perform abstract inductive reasoning without relying on memorization or external knowledge. Even with careful prompting, GPT-4o correctly solved only 19% of the tasks, far below the ~75% average human performance [904, 905]. Zhu et al. [906] suggested a four-level classification of AI intelligence, including L1 (arbitrating isputes), L2 (auditing a review), L3 (reviewing a paper), and L4 (authoring a paper). They classify the current state-of-the-art LLM-driven agents as approaching L2-level capabilities. To enhance agents’ reasoning abilities, researchers have introduced techniques such as chain-of-thought [907], tree-of-thoughts [72], and [70]. Although new methods continue to emerge, as discussed in Section 2.2, further advancements in reasoning capacity remain crucial for achieving reliable causal inference in scientific research.

**Agents and LLMs also struggle with quantitative and symbolic problems.** For example, GPT-4 and GPT-3.5 often struggle with reliably performing complex arithmetic such as multiplying  $12,345 \times 98,765$ , or translating IUPAC chemical names into accurate molecular graphs [908, 697]. A common approach to overcoming these limitations is to use external tools rather than relying on the LLM itself for reasoning. In mathematical problem-solving, for example, tools like symbolic solvers are preferred over direct LLM inference [753]. However, this mitigation does not resolve the intrinsic deficiency in numerical understanding, which poses a potential risk to scientific reasoning. Moreover, Yu et al. [909] found that tool-augmented LLMs do not consistently outperform base LLMs without tools in chemistry problem-solving. For instance, for *specialized* chemistry tasks, such as synthesis prediction, augmenting

LLMs with specialized tools can boost the performance substantially; however, tool augmentation is less effective for *general* chemistry questions, such as those in exams, where no specific tools can directly solve a given question. In these scenarios, an agent’s ability to reason correctly by using multiple pieces of chemistry knowledge becomes more important.

The preceding discussion emphasizes the importance of developing robust methodologies for evaluating AI agents as scientific research assistants, a topic discussed at length by Cappello et al. [910].

### 12.3.3 Challenges in Integrating Prior Knowledge

Prior knowledge is a crucial factor for higher intelligence. As discussed in Section 12.1, the agent’s prior knowledge,  $M_t^{\text{mem}}$ , helps decrease  $D_K(\theta, M_t^{\text{mem}})$  and increase the agent’s intelligence,  $IQ_t^{\text{agent}}$ . Human-led scientific discoveries frequently achieve breakthroughs with relatively small datasets, thanks to the vast prior knowledge humans possess. The start-of-the-art LLMs that power autonomous agents are trained on nearly all publicly available textual data, including websites, books, and other sources, thereby encompassing most common knowledge as well as publicly accessible specialized knowledge. However, achieving an agent that can seamlessly integrate all existing human knowledge remains a significant challenge.

At least three types of knowledge sources may not be included in LLM pre-training: (1) Paywalled or unpublished knowledge, including non-open-access publications, industry-specific data, and failed experiments [911]. They are often not accessible to public models despite their potential value in refining domain-specific insights. (2) Empirical knowledge. Heuristic decisions by experts are often effective, particularly in scenarios where no existing data is available for a new problem. However, large amounts of expert heuristics are typically not accessible as textual data. (3) Contextual or situational knowledge. Knowledge related to real-world conditions, such as safety protocols in chemical reactions or equipment handling, is often absent from pre-trained models but is essential for practical applications.

Additionally, integrating diverse knowledge sources presents challenges in reconciling conflicting information. For example, OpenAI’s Deep Research [912] actively gathers online information and performs multi-step reasoning, achieving state-of-the-art performance on Humanity’s Last Exam and the GAIA benchmark. However, it still struggles to distinguish between authoritative information and rumors and exhibits limitations in confidence calibration, often misrepresenting its level of certainty [912]. Establishing a system to assess the levels of evidence [913] of different knowledge fragments—such as quantifying reliability and verifying references—may be necessary for effective knowledge fusion.